PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5 : G06G 7/00

A1

(11) International Publication Number: WO 91/18361

(43) International Publication Date: 28 November 1991 (28.11.91)

(21) International Application Number: PCT/US91/03504

(22) International Filing Date: 17 May 1991 (17.05.91)

(30) Priority data:
525,973 18 May 1990 (18.05.90) US

(71) Applicant: YALE UNIVERSITY [US/US]; 246 Church Street, Suite 401, New Haven, CT 06510 (US).

(72) Inventors: COIFMAN, Ronald ; 11 Hickory Hill Road, North Haven, CT 06473 (US). MEYER, Yves ; 49, boulevard Saint-Marcel, F-75013 Paris (US). WICKERHAUSER, Mladen, V. ; 18 Nash Street, New Haven, CT 06511 (US).

(74) Agent: NOVACK, Martin, M.; P.O. Box 901, Westport, CT 06881 (US).

(81) Designated States: AT (European patent), BE (European patent), CA, CH (European patent), DE (European patent), DK (European patent), ES (European patent), FR (European patent), GB (European patent), GR (European patent), IT (European patent), JP, LU (European patent), NL (European patent), SE (European patent).

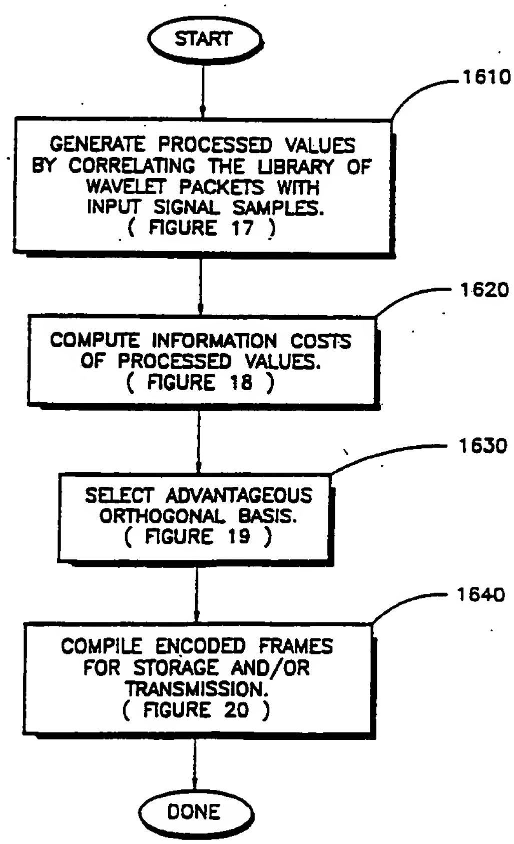
Published

With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: METHOD AND APPARATUS FOR ENCODING AND DECODING USING WAVELET-PACKETS

(57) Abstract

The disclosure involves the use of a library of modulated wavelet-packets which are effective in providing both precise frequency localization and space localization. An aspect of the disclosure involves feature extraction by determination of the correlations of a library of waveforms with the signal being processed, while maintaining orthogonality of the set of waveforms selected (i.e. a selected advantageous basis). In a disclosed embodiment, a method is provided for encoding and decoding an input signal, comprising the following steps (80): applying combinations of dilations and translations of a wavelet to the input signal to obtain processed values (1610); computing the information costs of the processed values (1620); selecting, as encoded signals, an orthogonal group of processed values, the selection being dependent on the computed information costs (1630); and decoding the encoded signals to obtain an output signal (1640). The wavelet preferably has a plurality of vanishing moments.



FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	ES	Spain	MG	Madagascar
AU	Australia	FI	Finland	ML	Mali
BB	Barbados	FR	France	MN	Mongolia
BE	Belgium	GA	Gabon	MR	Mauritania
BF	Burkina Faso	GB	United Kingdom	MW	Malawi
BG	Bulgaria	GN	Guinea	NL	Netherlands
BJ	Benin	GR	Greece	NO	Norway
BR	Brazil	HU	Hungary	PL	Poland
CA	Canada	IT	Italy	RO	Romania
CF	Central African Republic	JP	Japan	SD	Sudan
CG	Congo	KP	Democratic People's Republic of Korea	SE	Sweden
CH	Switzerland			SN	Senegal
CI	Côte d'Ivoire	KR	Republic of Korea	SU	Soviet Union
CM	Cameroon	LI	Liechtenstein	TD	Chad
CS	Czechoslovakia	LK	Sri Lanka	TG	Togo
DE	Germany	LU	Luxembourg	US	United States of America
DK	Denmark	MC	Monaco		

Description

METHOD AND APPARATUS FOR ENCODING AND DECODING USING WAVELET-PACKETS

FIELD OF THE INVENTION

This invention relates to a method and apparatus for encoding and decoding signals which may represent any continuous or discrete values.

BACKGROUND OF THE INVENTION

It is well established that various types of signals can be efficiently encoded and subsequently decoded in a manner which substantially reduces the size of the information required (e.g. number of bits, bandwidth, or memory) without undue or noticeable degradation of the decoded signal. Examples are the types of audio and video bandwidth compression schemes that are currently in widespread use.

In signal analysis, it is often useful to recognize the appearance of characteristic frequencies, but this knowledge generally has to be coupled with the location of the time (or space) interval giving rise to the frequency. Such questions can usually be tackled by use of the windowed Fourier transform, with different size windows corresponding to the scale of the transient feature. This analysis can be achieved by correlating the signal to all windowed exponentials and checking for large correlations. Unfortunately, due to lack of independence, information obtained can be redundant and inefficient for feature reconstruction purposes.

It is among the objects of the present invention to provide an improved encoding and decoding method and apparatus which overcomes limitations of prior art techniques and provides improved and more efficient operation.

SUMMARY OF THE INVENTION

The present invention involves, in part, the use of a library of modulated wavelet-packets which are effective in

providing both precise frequency localization and space localization. An aspect of the invention involves feature extraction by determination of the correlations of a library of waveforms with the signal being processed, while maintaining orthogonality of the set of waveforms selected (i.e. a selected advantageous basis).

In accordance with an embodiment of the invention, a method is provided for encoding and decoding an input signal, comprising the following steps: applying combinations of dilations and translations of a wavelet to the input signal to obtain processed values; computing the information costs of the processed values; selecting, as encoded signals, an orthogonal group of processed values, the selection being dependent on the computed information costs; and decoding the encoded signals to obtain an output signal. As used herein, wavelets are zero mean value orthogonal basis functions which are non-zero over a limited extent and are used to transform an operator by their application to the operator in a finite number of scales (dilations) and positions (translations) to obtain transform coefficients. [In the computational context, very small non-zero values may be treated as zero if they are known not to affect the desired accuracy of the solution to a problem.] A single averaging wavelet of unity mean is permitted. Reference can be made, for example, to: A. Haar, Zur Theorie der Orthogonalen Functionsysteme, Math Annal. 69 (1910); K.G. Beauchamp, Walsh Functions And Their Applications, Academic Press (1975); I. Daubechies, Orthonormal Bases of Compactly Supported Wavelets, Comm. Pure Appl. Math XL1 (1988).

In a preferred embodiment of the invention, the wavelet has a plurality of vanishing moments. In this embodiment, the step of applying combinations of dilations and translations of the wavelet to the input signal to obtain processed values comprises correlating said combinations of dilations and translations of the wavelet with the input signal. The combinations of dilations and translations of the wavelet are designated as wavelet-packets, and in a disclosed embodiment the step of applying wavelet-packets to the input signal to obtain processed

values includes: generating a tree of processed values, the tree having successive levels obtained by applying to the input signal, for a given level, wavelet-packets which are combinations of the wavelet-packets applied at a previous level. Also in a disclosed embodiment, the steps of computing information costs and selecting an orthogonal group of processed values includes performing said computing at a number of different levels of said tree, and performing said selecting from among the different levels of the tree to obtain an orthogonal group having a minimal information cost (the "best basis"). Also in this embodiment, the step of selecting an orthogonal group of processed values includes generating encoded signals which represent said processed values in conjunction with their respective locations in said tree.

Further features and advantages of the invention will become more readily apparent from the following detailed description when taken in conjunction with the accompanying drawings.

BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1 is a block diagram of an apparatus in accordance with an embodiment of the invention, and which can be used to practice the method of the invention.

Fig. 2 is a diagram illustrating Haar functions and combinations of dilations and translations of such functions.

Fig. 3 illustrates a tree of nodes.

Fig.s 4A, 4B and 4C illustrate examples of possible orthogonal bases.

Fig. 5 illustrates a wavelet having two vanishing moments.

Fig.s 6-13 illustrates examples of wavelet-packets.

Fig. 14 illustrates an example of how information cost can be computed.

Fig.s 15A and 15B illustrate a procedure for reconstruction in accordance with an embodiment of the invention.

Fig. 16 shows a flow diagram which is suitable for controlling the encoder processor to implement an embodiment of the encoding apparatus and method in accordance with the

invention.

Fig. 17 is a flow diagram of a routine for generating processed values from the sampled signal using a wavelet-packet basis.

Fig. 18 is a flow diagram of a routine for computing information cost.

Fig. 19 is a routine for selecting an advantageous or best orthogonal basis.

Fig. 20 is a flow diagram of a routine for generating output encoded words.

Fig. 21 is a flow diagram of the decoder routine for processing frames of words and reconstructing the orthogonal basis indicated by the words of a frame.

Fig. 22 is a further portion of the routine for reconstruction in the decoder.

DESCRIPTION OF THE PREFERRED EMBODIMENTS

Referring to Fig. 1, there is shown a block diagram of an apparatus in accordance with an embodiment of the invention for encoding and decoding an input signal which can be any continuous or discrete signal or sequence of numbers representing values in one or more dimensions (e.g. audio, still or moving pictures, atmospheric measurement data, etc.) and which, for purposes of illustration, can be considered as an audio signal $x(t)$. At the encoder 100 the signal is coupled to an analog-to-digital converter 102, which produces signal samples x_1, x_2, x_3..., a sequence of which can be envisioned as a vector x. The digital samples are coupled to an encoder processor 105 which, when programmed in the manner to be described, can be used to implement an embodiment of the invention and to practice an embodiment of the method of the invention. The processor 105 may be any suitable processor, for example an electronic digital or analog processor or microprocessor. It will be understood that any general purpose or special purpose processor, or other machine or circuitry that can perform the computations described herein, electronically, optically, or by other means, can be utilized. The processor 105, which for purposes of the

particular described embodiments hereof can be considered as the processor or CPU of a general purpose electronic digital computer, such as a SUN-3/50 Computer sold by Sun Microsystems, will typically include memories 125, clock and timing circuitry 130, input/output functions 135 and display functions 140, which may all be of conventional types.

With the processor appropriately programmed, as described hereinbelow, a compressed output signal x^c is produced which is an encoded version of the input signal, but which requires less bandwidth. In the illustration of Fig. 1, the encoded signals x^c are shown as being coupled to a transmitter 140 for transmission over a communications medium (air, cable, fiber optical link, microwave link, etc.) 150 to a receiver 160. The encoded signals are also illustrated as being coupled to a storage medium 142, which may alternatively be part of the processor subsystem 105, and are also illustrated as being manipulated such as by multiplication by a sparse matrix M^{wp}, as described in the U.S. Patent Application Serial No. 525,974. [See also Appendix III.] The matrix M^{wp} can obtain be obtained using the wavelet-packet best basis selection hereof (see also Appendix V). The signal itself may, of course, also be in the form of a matrix (i.e., a collection of vectors). The stored and/or manipulated signals can be decoded by the same processor subsystem 105 (suitably programmed, as will be described) or other decoding means.

In the illustrated embodiment, another processor 175, which is shown as being similar to the processor 105, also includes memories 225, clock and timing circuitry 230, input/output functions 235, and display functions 240, which may again be of conventional types. Processor 175 is employed, when suitably programmed as described, to decode the received encoded signal x^c, and to produce an output digital signal $x_1', x_2', x_3', \ldots$, (or vector x') which is a representation of the input digital signal, and which can be converted, such as by digital-to-analog converter 195, to obtain an analog representation $x'(t)$ of the original input signal $x(t)$. As will become understood, the accuracy of the representation will depend upon the encoding process and the degree of bandwidth compression.

Before describing the underlying theory of the invention, it is noted that reference can be made to Appendices I-V, appended hereto, for supplemental description of the theoretical foundations and further approaches.

A well known wavelet basis, which has a single vanishing moment, as defined hereinbelow, is the Haar basis [see, for example, A. Haar, Zur Theorie Der Orthogonalen Funktionensysteme, Math Annal. 69, 1910, and Appendix I of the abovereferenced U.S. Patent Application Serial No. 525,974]. Consider the Haar basis as applied to a simplified case of eight samples x_1, $x_2 \ldots x_8$. For uniform amplitudes, and ignoring normalizing coefficients (which are multiples of $1/\sqrt{2}$ for Haar wavelets), a set of waveforms can be developed from combinations of Haar wavelets, as illustrated in Fig. 2 and in accordance with the following relationships:

$$\begin{aligned} s_1 &= x_1 + x_2 \\ s_2 &= x_3 + x_4 \\ s_3 &= x_5 + x_6 \\ s_4 &= x_7 + x_8 \\ d_1 &= x_1 - x_2 \\ d_2 &= x_3 - x_4 \\ d_3 &= x_5 - x_6 \\ d_4 &= x_7 - x_8 \end{aligned} \tag{1}$$

$$\begin{aligned} ss_1 &= s_1 + s_2 \\ ss_2 &= s_3 + s_4 \\ ds_1 &= s_1 - s_2 \\ ds_2 &= s_3 - s_4 \\ sd_1 &= d_1 + d_2 \\ sd_2 &= d_3 + d_4 \\ dd_1 &= d_1 - d_2 \\ dd_2 &= d_3 - d_4 \end{aligned} \tag{2}$$

$$\begin{aligned} sss_1 &= ss_1 + ss_2 \\ dss_1 &= ss_1 - ss_2 \\ sds_1 &= ds_1 + ds_2 \\ dds_1 &= ds_1 - ds_2 \end{aligned}$$

$$\begin{aligned} ssd_1 &= sd_1 + sd_2 \\ dsd_1 &= sd_1 - sd_2 \\ sdd_1 &= dd_1 + dd_2 \\ ddd_1 &= dd_1 - dd_2 \end{aligned} \qquad (3)$$

The first group of relationships, (1) [the top eight waveforms in Fig. 2], are Haar functions, and are orthogonal. The last group of relationships, (3) [the bottom eight waveforms in Fig. 2], are the first eight of the well known Walsh functions. As is known in the art, the Walsh functions are orthogonal and complete, and can be advantageously used to transform and subsequently back-transform certain types of signals to achieve, inter alia, signal compression. It can be observed, however, that the set of the entire twenty-four functions of Fig. 2 is not orthogonal, which follows from the fact that some of the functions are derived from combinations of other functions. For example, sd_1 is the sum of d_1 and d_2, and is not orthogonal to d_1 or to d_2.

The sums and differences of relationships (1)-(3) are arranged in a tree of "nodes" in Fig. 3. Four "levels" are shown, level 0 being the sample data, and levels 1, 2 and 3 respectively corresponding to the groups of relationships (1), (2) and (3) above. The boxes ("nodes") at each level contain respective sum and difference terms, and are connected by branches to the nodes from which they are derived. It is seen that level 1 has two nodes (labelled, from left to right, 1 and 2), level 2 has four nodes (labelled, from left to right, 1-4), and level 3 has eight nodes (labelled, from left to right, 1-8). It follows that a level k would have 2^k nodes. The "positions" of the functions (or members) within a node are also numbered, from left to right, as illustrated in node 1 of level 1 (only). The nodes of level 1 each have four positions, the nodes of the level 2 each have two positions, and the nodes of level 3 each have one position. It is seen that each "parent" node has two "child" nodes, the members of the children being derived from those of the parent. Thus, for example, node 1 of level 2 and node 2 of level 2 are both children of node 1 of level 1. This follows from the relationships (2) set forth above. In particular, the members of node 1, level 2 (ss_1 and ss_2) are

derived from sums of members of node 1, level 1, and the members of node 2, level 2 (ds_1 and ds_2) are derived from differences of members of node 1, level 1.

In accordance with an aspect of the present invention, a complete set of functions (that is, a set or "basis" which permits reconstruction of the original signal samples) is obtained from the tree, permitting selection of nodes from any level. The selection is made in a manner which minimizes the information cost of the basis; i.e., the selected basis can be represented with a minimum bandwidth requirement or minimum number of bits for a given quality of information conveyed. Orthogonality of the selected basis is maintained by following the rule that no ancestor (parent, grandparent, etc.) of a selected node is used. [Conversely, no descendant (child, grandchild, etc.) of a selected node is used.] For the simplified tree of Fig. 3, there are twenty five possible orthogonal basis selections. Using shaded boxes to indicate the nodes selected for a given basis, four examples of possible orthogonal bases are shown in Figs. 4A-4D. If desired a basis which is the best level basis could alternatively be determined and used.

The Haar wavelet system, and wavelet-packets derived therefrom, has been used so far for ease of explanation. In accordance with an aspect of the present invention, advantageous wavelet-packets are generated using wavelets having a plurality of vanishing moments and, preferably, several vanishing moments. For description of the types of wavelets from which these wavelet-packets can be derived, reference can be made to I. Daubechies, Orthonormal Bases of Compactly Supported Wavelets, Comm. Pure, Applied Math, XL1, 1988; Y. Meyer Principe d'Incertitude, Bases Hilbertiennes et Algèbres d'Opèrateurs, Sèminaire Bourbaki, 1985-86, 662, Astèrisque (Sociète' Mathèmatique de France); S. Mallat, Review of Multifrequency Channel Decomposition of Images and Wavelet Models, Technical Report 412, Robotics Report 178, NYU (1988), and the abovereferenced U.S. Patent Application Serial No. 525,974.

The wavelet illustrated in Fig. 5 has two vanishing

moments. As used herein, the number of vanishing moments, for a wavelet $\Psi(x)$ is determined by the highest integer n for which

$$\int \Psi(x)\ x^k dx = 0$$

where $0 \leq k \leq n-1$, and this is known as the vanishing moments condition. Using this convention, the Haar wavelet has 1 vanishing moment, and the wavelet of Fig. 5 has 2 vanishing moments. [It is generally advantageous to utilize wavelets having as many vanishing moments as is practical, it being understood that the computational burden increases as the number of vanishing moments (and coefficients) increases. Accordingly, a trade-off exists which will generally lead to use of a moderate number of vanishing moments.] The wavelet of Fig. 5 has coefficients as follows (see e.g. Daubechies, supra):

$$h_1 = (1 + \sqrt{3})/4\sqrt{2}$$
$$h_2 = (3 + \sqrt{3})/4\sqrt{2}$$
$$h_3 = (3 - \sqrt{3})/4\sqrt{2}$$
$$h_4 = (1 - \sqrt{3})/4\sqrt{2}$$
$$g_1 = h_4$$
$$g_2 = -h_3$$
$$g_3 = h_2$$
$$g_4 = -h_1$$

The vanishing moments condition, written in terms of defining coefficients, would be the following:

$$\sum_{i=1}^{L} g_i i^k = 0$$

where

$$0 \leq k \leq n-1 \quad \text{and}$$
$$g_i = (-1)^{L-i+1} h_{L-i+1}$$

and L is the number of coefficients.

The procedure for applying this wavelet is similar to that for the Haar wavelet, but groups of four elements are utilized and are multiplied by the h coefficients and the g coefficients to obtain the respective terms of opposing polarity. To obtain wavelet-packets from the wavelets, the previously illustrated levels of sum and difference terms are obtained using the h and g

coefficients, respectively. The "sum" correlation terms for the first level are computed as follows:

$$\begin{aligned} s_1 &= h_1x_1 + h_2x_2 + h_3x_3 + h_4x_4 \\ s_2 &= h_1x_3 + h_2x_4 + h_3x_5 + h_4x_6 \\ s_3 &= h_1x_5 + h_2x_6 + h_3x_7 + h_4x_8 \\ &\cdot \\ &\cdot \\ &\cdot \\ s_{k/2} &= h_1x_{k-1} + h_2x_k + h_3x_{k+1} + h_4x_{k+2} \end{aligned} \tag{4}$$

The "difference" correlation terms for the first level are computed as follows:

$$\begin{aligned} d_1 &= g_1x_1 + g_2x_2 + g_3x_3 + g_4x_4 \\ d_2 &= g_1x_3 + g_2x_4 + g_3x_5 + g_4x_6 \\ d_3 &= g_1x_5 + g_2x_6 + g_3x_7 + g_4x_8 \\ &\cdot \\ &\cdot \\ &\cdot \\ d_{k/2} &= g_1x_{k-1} + g_2x_k + g_3x_{k+1} + g_4x_{k+2} \end{aligned} \tag{5}$$

The four sets of second level sum and difference correlation terms can then be computed from the first level values as follows:

$$\begin{aligned} ss_1 &= h_1s_1 + h_2s_2 + h_3s_3 + h_4s_4 \\ ss_2 &= h_1s_3 + h_2s_4 + h_3s_5 + h_4s_6 \\ ss_3 &= h_1s_5 + h_2s_6 + h_3s_7 + h_4s_8 \\ &\cdot \\ &\cdot \\ &\cdot \\ ss_{k/4} &= h_1s_{k/2-1} + h_2s_{k/2} + h_3s_{k/2+1} + h_4s_{k/2+2} \end{aligned} \tag{6}$$

$$\begin{aligned} ds_1 &= g_1s_1 + g_2s_2 + g_3s_3 + g_4s_4 \\ ds_2 &= g_1s_3 + g_2s_4 + g_3s_5 + g_4s_6 \\ ds_3 &= g_1s_5 + g_2s_6 + g_3s_7 + g_4s_8 \\ &\cdot \\ &\cdot \\ &\cdot \\ ds_{k/4} &= g_1s_{k/2-1} + g_2s_{k/2} + g_3s_{k/2+1} + g_4s_{k/2+2} \end{aligned} \tag{7}$$

$$
\begin{aligned}
sd_1 &= h_1d_1 + h_2d_2 + h_3d_3 + h_4d_4 \\
sd_2 &= h_1d_3 + h_2d_4 + h_3d_5 + h_4d_6 \\
sd_3 &= h_1d_5 + h_2d_6 + h_3d_7 + h_4d_8 \\
&\cdot \\
&\cdot \\
&\cdot \\
sd_{k/4} &= h_1d_{k/2-1} + h_2d_{k/2} + h_3d_{k/2+1} + h_3d_{k/4+2}
\end{aligned}
\tag{8}
$$

$$
\begin{aligned}
dd_1 &= g_1d_1 + g_2d_2 + g_3d_3 + g_4d_4 \\
dd_2 &= g_1d_3 + g_2d_4 + g_3d_5 + g_4d_6 \\
dd_3 &= g_1d_5 + g_2d_6 + g_3d_7 + g_4d_8 \\
&\cdot \\
&\cdot \\
&\cdot \\
dd_{k/4} &= g_1d_{k/2-1} + g_2d_{k/2} + g_3d_{k/2+1} + g_4d_{k/2+2}
\end{aligned}
\tag{9}
$$

, and so on. Extra values at end positions can be handled by "wrap-around", truncation, or other known means of handling end conditions. It will be understood that the procedure described in conjunction with relationships (4)-(9) is operative to successively correlate the signal samples with the wavelet-packets for each successive level. If desired, the correlations can be implemented by generating wavelet-packets a priori (using the indicated coefficients, for this example), and then individually correlating wavelet-packets with the signal using analog (e.g. electronic or optical means), digital, or any suitable technique. If desired, a special purpose network could be used to perform the correlations.

In terms of the diagram of Fig. 3, the sums (4) and the differences (5) would occupy nodes 1 and 2, respectively, of level 1, and the sums (6), differences (7), sums (8) and differences (9) would occupy the nodes 1, 2, 3 and 4, respectively, of level 2.

It will be understood that in many practical situations the number of samples considered in a frame or window (level 0) will be larger than 8, and the tree will also be larger than those shown here for ease of illustration. Fig.s 6-10 show the first

five wavelet-packets synthesized for sample length 1024, using a wavelet with six coefficients (six h's and six g's), and Fig.s 11-13 illustrate three of the higher frequency wavelet-packets.

In accordance with an aspect of the invention, the basis is selected in a manner which minimizes information cost. There are various suitable ways in which information cost can be computed. Fig. 14 shows a parent node and its two children nodes. As a measure of information, one can compute the number N_p of correlation values in the parent node that exceed the particular threshold and the total number N_c of correlation values in the child nodes that exceed the particular threshold. As represented in Fig. 14, if N_p is less than or equal to N_c, the parent node will be preferred, whereas if N_p is greater than N_c, the children nodes will be preferred. As higher level comparisons are made (with the ancestors of the parent) the selection may be supplanted by an ancestor.

Another measure of information cost that can be used is the entropy cost function that is well known in information theory, and is threshold independent (see Appendix IV). Suitable weighting of coefficients can also be used. For example, if it is known or computed that certain values should be filtered, emphasized, or ignored, weighting can be appropriately applied for this purpose.

Fig. 15A illustrates a procedure for reconstruction which can be utilized at the decoder processor. The shaded boxes indicate the nodes defining the orthogonal basis that was selected at the encoder and is to be decoded. The arrows illustrate the reconstruction paths, and the cross-hatched boxes indicate reconstructed nodes, the last (level 0) reconstruction providing the reconstructed decoder output information. In particular, node 1, level 2 is reconstructed from node 1, level 3 and node 2, level 3. Node 1, level 1 is then reconstructed from node 1, level 2 and node 2, level 2, and so on.

Fig. 15B shows children nodes containing sy_1, sy_2, ...sy_n and dy_1, dy_2, ...dy_n being mapped into their parent node to obtain the reconstructed y_1, y_2, ...y_{2n}. If the four coefficients h_1, h_2, h_3 and h_4 (with the corresponding g_1, g_2, g_3 and g_4) were used for

encoding, the decoding relationships will be as follows:

<u>For y odd</u>

$$\begin{aligned} y_1 &= h_1 sy_1 + h_3 sy_0 + g_1 dy_1 + g_3 dy_0 \\ y_3 &= h_1 sy_2 + h_3 sy_1 + g_1 dy_2 + g_3 dy_1 \\ &\cdot \\ &\cdot \\ &\cdot \\ y_{2n+1} &= h_1 sy_{n+1} + h_3 sy_n + g_1 dy_{n+1} + g_3 dy_n \end{aligned} \tag{10}$$

<u>For y even</u>

$$\begin{aligned} y_2 &= h_2 sy_1 + h_4 sy_0 + g_2 dy_1 + g_4 dy_0 \\ y_4 &= h_2 sy_2 + h_4 sy_1 + g_2 dy_2 + g_4 dy_1 \\ &\cdot \\ &\cdot \\ &\cdot \\ y_{2k} &= h_2 sy_n + h_4 sy_{n-1} + g_2 dy_n + g_4 dy_{n-1} \end{aligned} \tag{11}$$

The values mapped into the parent mode are accumulated and, as in the encoder, extra values at the end positions (e.g. at y_0 above) can be handled by "wrap-around", truncation, or other known means of handling end conditions.

Referring to Fig. 16, there is shown a flow diagram which, when taken together with the further flow diagrams referred to therein, is suitable for controlling the processor to implement an embodiment of the encoding apparatus and method in accordance with the invention. The block 1610 represents the generating of processed values by correlating wavelet-packets with the input signal samples. There are various ways in which this can be achieved, the routine of Fig. 17 describing an implementation of the present embodiment. The block 1620 represents the routine, described further in conjunction with Fig. 18, for computing the information costs of the processed values. As described further hereinbelow, there are various ways of computing the measure or cost of information contained in the processed values. In an illustrated embodiment, a thresholding procedure is utilized, and information cost is determined by the number of values which exceed a particular threshold. The block 1630 represents the routine of Fig. 19 for selection of an advantageous orthogonal

basis from among the processed values, the selection of the basis being dependent on the computed information costs. The block 1640 represents compiling encoded frames from the selected processed values which constitute the basis, such as for subsequent recovery after processing, storage, and/or transmission. This routine is described in conjunction with Fig. 20.

Referring to Fig. 17, there is shown a flow diagram of a routine for generating the processed values from the sampled signal, or signal portion, using a wavelet-packet basis. The block 1710 represents the reading in of the selected coefficients h_i, g_i. The block 1720 is then entered, this block representing the initializing of a level index at 0, the initializing of a node index at 1, and the initializing of a position index at 1. The sample data, considered as level 0 is then read in, as represented by the block 1725. The sample data may consist, for example, of 256 sequential samples of an acoustical signal to be compressed and transmitted. The level index is then incremented, as represented by block 1730, and the first level processed values are computed and stored in accordance with the relationships (4) and (5) set forth above (block 1735 and loops 1748 and 1760). For example, for the first position of the first node of level 1, s_1 will be computed. If the wavelet employed is representable by a filter having four coefficients, as in the example above, s_1 will be computed as the sum of h_1x_1, h_2x_2, h_3x_3, h_4x_4. If a wavelet of more vanishing moments is used, more coefficients will be employed. In general, it will be preferable to utilize a wavelet having several coefficients, greater than four, the above examples being set forth for ease of illustration.

In loop 1748, inquiry is made (diamond 1740) as to whether the last position of the current node has been reached. If not, the position index is incremented (block 1745), and the block 1735 is re-entered for computation of the next processed value of the current node and level. The loop 1748 is then continued until all processed values have been computed for the current node, whereupon inquiry is made (diamond 1750) as to whether the

last node of the current level has been reached. If not, the node index is incremented (block 1755) and the loop 1760 is continued until the processed values have been computed for all nodes of the current level. For the first level, there will be only two nodes, with the values thereof being computed in accordance with the relationships (4) and (5) set forth above.

When the inquiry of diamond 1750 is answered in the affirmative, diamond 1770 is entered, and inquiry is made as to whether the last level has been processed. If not, the block 1730 is re-entered, to increment the level index, and the loop 1780 is continued until processed values have been obtained for the nodes at all levels of the tree.

Referring to Fig. 18, there is shown a flow diagram of the routine for computing the information cost of the nodes of the tree, so that an advantageous orthogonal basis can be selected. The block 1810 represents initializing the level index to the highest level (e.g., the last level in the illustration of Fig. 3). The node index and the position index are initialized at 1 (blocks 1815 and 1820). A node content count, which is used in the present embodiment to keep track of the number of processed values in a node that exceed a predetermined threshold, is initialized at zero, as represented by the block 1825. Inquiry is then made (diamond 1830) as to whether the value at the current position is less than a predetermined threshold value. If not, the node content count is incremented (block 1835), and the diamond 1840 is entered. If, however, the processed value at the current position is less than the threshold value, the diamond 1840 is entered directly. [At this point, the processed value could be set to zero prior to entry of diamond 1840 from the "yes" output branch of diamond 1830, but it is more efficient to handle this later.] Inquiry is then made (diamond 1840) as to whether the last position of the node has been reached. If not, the position index is incremented (block 1850), diamond 1830 is re-entered, and the loop 1855 is continued until all processed values of the current node have been considered. When this occurs, the node content count is stored for the current node (of the current level), as represented by the block 1860. Inquiry

is then made (diamond 1865) as to whether the last node of the level has been processed. If not, the block 1870 is entered, the node index is incremented, the block 1820 is re-entered, and the loop 1875 is continued until all nodes of the current level have been considered. Inquiry is then made (diamond 1880) as to whether the current level is the highest level. If so, there is no higher level against which comparison of parent-to-children node comparisons can be made. In such case, the level index is incremented (block 1885), block 1815 is re-entered, and the procedure just described is repeated to obtain and store node content counts for each node of the next-to-highest level. When this has been done, the inquiry of diamond 1880 will be answered in the negative, and the next routine (Fig. 19) will be operative to compare the level (which has just been processed to compute information cost of each node) with the children nodes of the previously processed higher level.

In particular, the level index is initialized (block 1905) to the highest level less one, and all nodes on the highest level are marked "kept". The node index is initialized (block 1910) and the node content count of the current node of the current level is compared (block 1920) to the sum of the node content counts of the two nodes which are children of the current node. [For example, if the current node is N_i and the current level is L_j, then the count for the current node is compared to the sum of the counts for nodes N_{2i-1} and N_{2i} of level L_{j+1}.] If the comparison shows that the parent has an equal or lower count, the parent is marked "kept", and the two children nodes are marked "not kept" (as represented by the block 1930). Conversely, if the comparison shows that the sum of two children nodes has a lower count than the parent node, each of the children nodes keeps its current mark, and the current parent node is marked "not kept" (as represented by the block 1940). In the case where the children nodes are preferred, the sum of the counts of the children nodes are attributed to the parent node (block 1945). By so doing, the lowest count will be used for subsequent comparisons as ancestors are examined. The attribution of the count to the parent node will not be problematic, since only

"kept" nodes will be considered in the next stage of processing. Inquiry is then made (diamond 1950) as to whether the last node of the current level has been reached. If not, the node index is incremented (block 1955), block 1920 is re-entered, and the loop 1958 is continued until all nodes at the current level have been considered. Inquiry is then made (diamond 1960) as to whether the current level is level 1. If not, the level index is decremented (block 1965), block 1815 (Fig. 18) is re-entered, and the loop 1970 continues until all levels have been considered. At this point, the nodes which define the basis to be used have been marked "kept" [possibly together with some of their descendent nodes], and correspond, for example, to the shaded nodes of the Fig. 4 illustrations.

Referring to Fig. 20, there is shown a flow diagram of a routine for generating output encoded words which, in the present embodiment, are collected in a frame which represents the encoded form of the data x_1, x_2, x_3 ... x_n. For example, for an acoustical signal, the frame may represent a particular number of acoustical samples, for example 256 samples. As a further example, for a video signal, the frame may represent a frame of video, portion thereof, or transformed frequency components thereof. The number of encoded words in a frame will generally vary as the information being encoded varies, and will also generally depend upon the level of the threshold employed, if a threshold is employed as in the present embodiment. Fig. 20 illustrates an embodiment of a routine for generating a frame of words for the basis that was selected using the routines of Fig.s 18 and 19. A tree location index will be calculated which points to nodes in the tree in depth-first order (or so-called "pre-order"), as is well known in the art. The tree location index is initialized to 1 at level 0, node 1 (block 2010). Inquiry is made (diamond 2015) as to whether the node at that tree location is market "kept", and, if not, diamond 2020 is entered directly, and inquiry is made as to whether the entire tree has been examined, as indicated by the tree location index. If the entire tree has been searched, block 2080 is entered and a "frame complete" indication can be generated. If not, then loop

2011 is continued until a node marked "kept" is encountered, or until the entire tree has been searched. If a node marked "kept" is encountered, block 2030 is entered, and the tree location index of this "kept" node is recorded in memory; suppose for example that it is called "X". The position index in the node is initialized (block 2035). Inquiry is then made (diamond 2040) as to whether the value at the current position is above the predetermined threshold. If not, diamond 2055 is entered directly, and no word is generated for the value at the current position. If the value is above the threshold, block 2045 is entered, this block representing the generation of a word which includes the current level, node, and position, and the value at the position. The block 2050 is then entered, this block representing the loading of the just generated word into an output register. Inquiry is then made (diamond 2055) as to whether the last position in the current node has been reached. If not, the position is incremented (block 2060), diamond 2040 is re-entered, and the loop 2061 is continued until all positions in the node "X" have been considered. It will be understood that various formats can be used to represent the words. For example, a specific number of bits can be used for each of the level, node, position, and value. Alternatively, words could be of different length, e.g. depending on information content or entropy, with suitable delineation between words, as is known in the art. Also, if desired, all words in a particular node could be encoded with a single indication of level and node, with individual indications of position-value pairs. Inquiry is next made (diamond 2070), as to whether the last node location in the tree in depth-first order has been reached. If not, the tree location index is incremented (block 2070), and inquiry is made as to whether the new node is a descendant of "X", by a comparison of depth-first indices well known in the art. When this is the case, diamond 2065 is re-entered, and the loop 2071 is continued until the first node which is not a descendant of "X" is encountered, or until there are no more nodes. When a first non-descendant of "X" is encountered, diamond 2015 is re-entered and the loop 2081 is continued until all nodes which

are both marked "kept" and have no ancestors marked "kept" have contributed to the frame. Such nodes contain a complete orthogonal group of wavelet-packet correlations (see also Appendix I, II, and V). When either loop 2011 or the loop 2071 terminates by exhaustion of the nodes, block 2080 is entered and a "frame complete" indication can be generated. If desired, the frame can then be read out of the encoder register. However, it will be understood that the encoder register can serve as a buffer from which words can be read-out synchronously or asynchronously, depending on the application.

Referring to Fig. 21, there is shown a flow diagram of the decoder routine for processing frames of words and reconstructing the orthogonal basis indicated by the words of a frame. The block 2110 represents the reading in of the next frame. In the described embodiment, it is assumed that the frames of words are read into a buffer (e.g. associated with decoder processor subsystem 170 of Fig. 1), and the individual words processed sequentially by placement into appropriate addresses (which can be visualized as the selected basis nodes of a tree - as in Fig. 4), from which reconstruction is implemented in the manner to be described. However, it will be understood that individual words can be received synchronously or asynchronously, or could be output in parallel into respective tree nodes, if desired. Also, as was the case with the encoder, parallel processing or network processing could be employed to implement reconstruction, consistent with the principles hereof. In the routine of Fig. 21, the next word of the frame is read (block 2115), and a determination is made as to whether the node and level of the word is occurring for the first time (diamond 2117). If so the node (and its level) is added to the list of nodes (block 2118). The value indicated in the word is stored (block 2120) at a memory location indicated by the level, node, and position specified in the word. It will be understood that memory need be allocated only for positions within the nodes designated by the read-in words. Inquiry is then made (diamond 2130) as to whether the last word of the frame has been reached. If not, the block 2115 is re-entered, and the loop 2135 is

continued until all words of the frame have been stored at appropriate locations. It will be understood that, if desired, the word locations (level, node, and position) could alternatively be stored, and the values subsequently recovered by pointing back to their original storage locations.

During the next portion of the decoder routine, as shown in Fig. 22, the values in the nodes on the list are utilized to implement reconstruction as in the diagram of Fig.s 15A and 15B, with parent nodes being reconstructed from children nodes until the level zero information has been reconstructed. During this procedure, when a parent node is reconstructed from its children nodes, the parent node is added to the list of nodes, so that it will be used for subsequent reconstruction. This part of the procedure begins by initializing to the first node on the list (block 2210). Next, the block 2215 represents going to the memory location of the node and initializing to the first position in the node. Inquiry is then made (diamond 2220) as to whether there is a non-zero value at the position. If not, diamond 2240 is entered directly. If so, the value at the position is mapped into the positions of the parent node, with accumulation, as described above in conjunction with relationships (10) and (11). Inquiry is then made (diamond 2240) as to whether the last position of the node has been reached. If not, the next position in the node is considered (block 2245), diamond 2220 is re-entered, and the loop 2250 continues until all positions in the node have been considered. It will be understood that, if desired, a marker or vector can be used to indicate strings of blank positions in a node, or to point only to occupied positions, so that a relatively sparse node will be efficiently processed. In this regard, reference can be made to the abovereferenced U.S. Patent Application Serial No. 525,974. When the last position of the node has been considered, the node is removed from the list of nodes, as represented by block 2255, and inquiry is made (diamond 2260) as to whether the parent node is at level 0. If so, diamond 2270 is entered directly. If, however, the parent node is not at level 0, the parent node is added to the list of nodes (block 2265). Inquiry is then made

(diamond 2270) as to whether the last node on the list has been reached. If not, the next node on the list is considered (block 2275), block 2215 is re-entered, and the loop 2280 is continued until processing is complete and the reconstructed values have been obtained. The decoder output values can then be read out (block 2290).

It will be understood that similar techniques can be employed at higher dimensions and in other forms (see e.g. Appendix V). More complicated tree structures, such as where a node has more than two children (e.g. Appendices II and V) can also be utilized.

The invention has been described with reference to particular preferred embodiments, but variations within the spirit and scope of the invention will occur to those skilled in the art. For example, it will be recognized that the wavelet upon which the wavelet-packets are based can be changed as different parts of a signal are processed. Also, the samples can be processed as sliding windows instead of segments.

Appendix I

Construction of Wavelet-Packets

We introduce a new class of orthonormal bases of $L^2(\mathbb{R}^n)$ by constructing a "library" of modulated wave forms out of which various bases can be extracted. In particular, the wavelet basis, the walsh functions, and rapidly oscillating wave packet bases are obtained.

We'll use the notation and terminology of [D], whose results we shall assume.

§1. We are given an exact quadrature mirror filter $h(n)$ satisfying the conditions of Theorem (3.6) in [D], p. 964, i.e.

$$\sum h(n-2k)h(n-2\ell) = \delta_{k,\ell}\ ,\ \sum h(n) = \sqrt{2}.$$

We let $g_k = h_{k+1}(-1)^k$ and define the operations F_i on $\ell^2(\mathbb{Z})$ into "$\ell^2(2\mathbb{Z})$"

(1.0)
$$\begin{aligned} F_0\{s_k\}(i) &= 2\sum s_k h_{k-2i} \\ F_1\{s_k\}(i) &= 2\sum s_k g_{k-2i}. \end{aligned}$$

The map $\mathbf{F}(s_k) = F_0(s_k) \oplus F_1(s_k) \in \ell^2(2\mathbb{Z}) \oplus \ell^2(2\mathbb{Z})$ is orthogonal and

(1.1)
$$F_0^* F_0 + F_1^* F_1 = I$$

We now define the following sequence of functions.

(1.2)
$$\begin{cases} W_{2n}(x) = \sqrt{2}\sum h_k W_n(2x-k) \\ W_{2n+1}(x) = \sqrt{2}\sum g_k W_n(2x-k) \end{cases}$$

Clearly the function $W_0(x)$ can be identified with the function φ in [D] and W_1 with the function ψ.

Let us define $m_0(\xi) = \frac{1}{\sqrt{2}}\sum h_k e^{-ik\xi}$ and

$$m_1(\xi) = -e^{i\xi}\bar{m}_0(\xi+\pi) = \frac{1}{\sqrt{2}}\sum g_k e^{ik\xi}$$

Remark: The quadrature mirror condition on the operation $\mathbf{F} = (F_0, F_1)$ is equivalent to the unitarity of the matrix

$$\mathcal{M} = \begin{bmatrix} m_0(\xi), m_1(\xi) \\ m_0(\xi+\pi), m_1(\xi+\pi) \end{bmatrix}$$

Taking Fourier transform of (1.2) when $n = 0$ we get

$$\hat{W}_0(\xi) = m_0(\xi/2)\hat{W}_0(\xi/2)$$

i.e.,

$$\hat{W}_0(\xi) = \prod_{j=1}^{\infty} m_0(\xi/2^j)$$

and

$$\hat{W}_1(\xi) = m_1(\xi/2)\hat{W}_0(\xi/2) = m_1(\xi/2)m_0(\xi/4)m_0(\xi/2^3)\cdots$$

More generally, the relations (1.2) are equivalent to

(1.3) $$\hat{W}_n(\xi) = \prod_{j=1}^{\infty} m_{\varepsilon_j}(\xi/2^j)$$

and $n = \sum_{j=1}^{\infty} \varepsilon_j 2^{j-1}$ ($\varepsilon_j = 0$ or 1).

We can rewrite (1.1) as follows.

(1.4) $$\begin{aligned} W_{2n}(x-\ell) &= \sqrt{2}\sum h_{j-2\ell}W_n(2x-j) = F_0\{W_n(2x-j)\}(\ell) \\ W_{2n+1}(x-\ell) &= \sqrt{2}\sum g_{j-2\ell}W_n(2x-j) = F_1\{W_n(2x-j)\}(\ell) \end{aligned}$$

where $W_n(2x-j)$ is viewed as a sequence in j for (x, n) fixed. Using (1.1) we find:

(1.5) $$\boxed{W_n(x-j) = \sqrt{2}\sum_i h_{j-2i}W_{2n}\left(\frac{x}{2}-i\right) + g_{j-2i}W_{2n+1}\left(\frac{x}{2}-i\right)}$$

In the case $n = 0$ we obtain:

(1.6) $$W_0(x-k) = \sqrt{2}\sum h_{k-2i}W_0\left(\frac{x}{2}-i\right) + g_{k-2i}W_1\left(\frac{x}{2}-i\right)$$

from which we deduce the usual decomposition of a function f in the space Ω_0 (V_0 in [D]) i.e., a function f of the form

$$
\begin{aligned}
f(x) &= \sum s_k^0 W_0(x-k) \\
&= \sqrt{2}\sum\left(\sum s_k^0 h_{k-2i}\right)W_0\left(\frac{x}{2}-i\right)+\sqrt{2}\sum\left(\sum s_k^0 g_{k-2i}\right)W_1\left(\frac{x}{2}-i\right) \\
&= \sum \frac{1}{\sqrt{2}}F_0(s_k^0)(i)W_0\left(\frac{x}{2}-i\right)+\sum\frac{1}{\sqrt{2}}F_1(s_k^0)(i)W_1\left(\frac{x}{2}-i\right)
\end{aligned}
$$

More generally, if we define

$$\Omega_n = \{f : f = \sum \omega_k W_n(x-k)\}. \tag{1.7}$$

We find

$$f(x) = \sum \frac{1}{\sqrt{2}}F_0(\omega_k)(i)W_{2n}\left(\frac{x}{2}-i\right)+\sum\frac{1}{\sqrt{2}}F_1(\omega_k)(i)W_{2n+1}\left(\frac{x}{2}-i\right) \tag{1.8}$$

or

$$\sqrt{2}f(2x) = h + g \qquad h \in \Omega_{2n}\ g \in \Omega_{2n+1}$$

We now prove

Theorem (1.1). *The functions $W_n(x-k)$ form an orthonormal basis of $L^2(\mathrm{R})$.*

Proof. We proceed by induction on n, assuming that $W_n(x-k)$ form an orthonormal set of functions and, proving that, $W_{2n}(x-k), W_{2n+1}(x-k)$ form an *o.n* set.

By assumption $\|\sqrt{2}f(2x)\|_2^2 = \sum \omega_k^2$ if $f \in \Omega_n$ from the quadrature mirror condition on (F_0, F_1) we get

$$\sum \omega_k^2 = \sum F_0(\omega_k)(i)^2 + F_1(\omega_k)(i)^2.$$

Since $F_0(\omega_k)(i) = \mu_i$, $F_0(\omega_k)(i) = \nu_i$ can be chosen as two arbitrary sequences of ℓ^2 (arising from $\omega = F_0^*\mu_i + F_1^*\nu_i$) it follows that

$$\int |\sum \mu_i W_{2n}(x-i) + \sum \nu_i W_{2n+1}(x-i)|^2 = \sum \mu_i^2 + \sum \nu_i^2$$

which is equivalent to $W_{2n}(x-i), W_{2n+1}(x-j)$ being an *o.n* set of functions.

Let us now define $\delta f = \sqrt{2}f(2x)$. Formula (1.8) shows that $\delta\Omega_n = \Omega_{2n} \oplus \Omega_{2n+1}$ as an orthogonal sum or,

$$
\begin{aligned}
\delta\Omega_0 - \Omega_0 &= \Omega_1 \\
\delta^2\Omega_0 - \delta\Omega_0 &= \delta\Omega_1 = \Omega_2 \oplus \Omega_3 \\
\delta^3\Omega_0 - \delta^2\Omega_0 &= \delta\Omega_2 \oplus \delta\Omega_3 = \Omega_4 \oplus \Omega_5 \oplus \Omega_6 \oplus \Omega_7 \text{ or} \\
\delta^k\Omega_0 - \delta^{k-1}\Omega_0 &= \Omega_{2^{k-1}} \oplus \Omega_{2^{k-1}+1} \cdots \oplus \Omega_{2^k-1}
\end{aligned} \tag{1.9}
$$

and

$$\delta^k\Omega_0 = \Omega_0 \oplus \Omega_1 \oplus \cdots \oplus \Omega_{2^k-1}$$

More generally, we let $\mathcal{W}_k = \delta^{k+1}\Omega_0 - \delta^k\Omega_0 = \delta^k\Omega_1 = \delta^k\mathcal{W}_1$. Therefore we have

Proposition (1.1).

$$\mathcal{W}_k = \delta^k W_1 = \Omega_{2^k} \oplus \Omega_{2^k+1} \oplus \cdots \oplus \Omega_{2^{k+1}-1}.$$

Alternatively, the functions

$$W_n(x-j) \qquad j \in \mathbf{Z}\ 2^k \leq n < 2^{k+1}$$

form an orthonormal basis of $\mathcal{W}_k$.

Since the spaces $\mathcal{W}_k$ are mutually orthogonal and span $L^2(\mathbf{R})$ see [D], it follows that $W_n(x-j)$ are complete.

§2. Orthonormal bases extracted out of the "library" $2^{j/2}W_n(2^jx-k)$.

We start by observing that $2^{j/2}W_1(2^jx-k)$ form a basis of $\mathcal{W}_j$ as we vary k and a basis of $L^2(\mathbf{R})$ as j,k vary. This is the wavelet basis constructed in [D]. The following simple generalization is useful to obtain a better localization in frequency space.

Theorem (2.1). *The functions*

$$2^{j/2}W_n(2^jx-k) \qquad j = 0, \pm1, \ldots, k = 0, \pm1, \pm2$$

$2^\ell \leq n < 2^{\ell+1}$ *for fixed* ℓ *form an orthonormal basis of* $L^2(\mathbf{R})$ $\ell = 0, 1, 2, \ldots$.

Remark: This is a wavelet basis with dyadic dilations and 2^ℓ fundamental wavelets.

Proof. We have seen, in Proposition (1.1), that $W_n(x-k)$ $2^\ell \leq n < 2^{\ell+1} k \in \mathbf{Z}$ form an *o.n* basis of $\mathcal{W}_\ell$, from which we deduce that $2^{j/2}W_n(2^jx-k)$ form an *o.n.* basis of $\mathcal{W}_{\ell+j}$ spanning $L^2(\mathbf{R})$ for $j \in \mathbf{Z}$ $k \in \mathbf{Z}$.

In the next example we vary j and n simultaneously to obtain a basis whose number of oscillation is inversely proportional to the length of its support.

Theorem (2.2). *Let* $\ell(n) = [\log_2 n]$ *i.e.,* $2^{\ell(n)} \leq n < 2^{\ell+1}$ *then*

$$W_n(2^{\ell(n)}x-k)2^{\ell(n)/2},\ W_n(2^{\ell(n)+1}x-k)2^{\frac{\ell(n)+1}{2}}$$

form an o.n basis of $L^2(\mathbf{R})$.

Proof. Fix $\ell(n) = \ell$, and consider n with $2^\ell \leq n < 2^{\ell+1}$. as seen in the proof of Theorem (1.2)

$$W_n(2^\ell x-k)2^{\ell/2} \quad \text{form an } o.n \text{ basis of } \mathcal{W}_{2\ell}$$

and

$$W_n(2^{\ell+1}x-k)2^{(\ell+1)/2} \quad \text{form an } o.n \text{ basis of } \mathcal{W}_{2\ell+1}$$

since $L^2 = \sum \oplus \mathcal{W}_{2\ell} \oplus \sum \oplus \mathcal{W}_{2\ell+1}$, we have a complete basis.

These can be generalized as follows.

Theorem (2.3). *Let a collection $\{\ell, n\}$ be given such that the dyadic intervals $I_{\ell,n} = [2^\ell n, 2^\ell(n+1))$ form a disjoint covering of $(0,\infty)$, then $2^{\ell/2}W_n(2^\ell x-k)$ form a complete orthonormal basis of $L^2(\mathbf{R})$.*

This theorem becomes obvious in the following case. Let

$$m_0(\xi) = \begin{cases} 1 & |\xi| < \frac{\pi}{2} \\ 0 & \frac{\pi}{2} \le |\xi| < \pi \end{cases}$$

a periodic function of period 2π, and $m_1(\xi) = 1 - m_0(\xi)$. Let

$$\hat{\omega}_n(\xi) = \prod_{j=1}^{\infty} m_{\varepsilon_j}(\xi/2^j) \quad n = \sum^{\infty} \varepsilon_j 2^{j-1}$$

then

$$\hat{\omega}_n(\xi) = \begin{cases} 1 & n \le |\xi/\pi| < n+1 \\ 0 & \text{elsewhere} \end{cases}$$

and the orthonormal basis $\omega_n(x-k)$ in Fourier space is

$$e^{ik\xi}\hat{\omega}_n(\xi)$$

which is the simplest variation on a "windowed" (2 windows) Fourier transform. Theorem (1.4) is obvious in this case. This theorem is also easy to understand from the point of view of subband coding as we shall see.

§3. Subband coding and expansions in terms of W_n

We assume, given a function which. on the scale 2^{-N}, is well approximated as

(3.1)
$$f(x) = \sum s_k^0 W_0(2^N x - k)2^{N/2}$$

as seen in (1.8)

(3.2)
$$\begin{aligned} f(x) &= \frac{1}{\sqrt{2}} \sum \left\{ F_0(s_k^0)(i) W_0\left(\frac{x2^N}{2} - i\right) + F_i(s_k^0)(i) W_1\left(\frac{x2^N}{2} - i\right) \right\} 2^{N/2} \\ &= \left\{ \frac{1}{\sqrt{2}} f_0\left(\frac{x}{2}2^N\right) + \frac{1}{\sqrt{2}} f_1\left(\frac{x}{2}2^N\right) \right\} 2^{N/2} \text{ with } f_0 \in \Omega_0 f_1 \in \Omega_1 \end{aligned}$$

The coefficients of f_0 are given by $F_0(s^0)$.
The coefficients of f_1 are given by $F_1(s^0)$.
Continuing an application of (1.8) gives

$$= \frac{2^{N/2}}{2}\left\{f_{00}\left(\frac{x}{4}2^N\right) + f_{10}\left(\frac{x}{4}2^N\right) + f_{01}\left(\frac{x}{4}2^N\right) + f_{11}\left(x22^N\right)\right\}$$

where f_{00}, f_{10} are obtained by decomposing f_0
and f_{01}, f_{11} are obtained by decomposing f_1
$f_{00} \in \Omega_0, f_{10} \in \Omega_1, f_{01} \in \Omega_2, f_{11} \in \Omega_3$. If we continue this decomposition and observe that the binary tree corresponds to the realization of n as $n = \sum_1 \varepsilon_j 2^{j-1}$ and that after ℓ iterations we get

(3.3) $$f(x) = 2^{(N-\ell)/2} \sum_{k=0}^{2^\ell - 1} f_n(x2^{N-\ell}) \text{ with } f_n \in \Omega_n.$$

and $f_n(x) = \sum_k \langle f, W_n(2^{N-\ell}x - k)\rangle W_n(2^{N-\ell}x - k)2^{N-\ell}$ with

(3.4) $$2\frac{N-\ell}{2}\langle f, W_n(2^{N-\ell} - k)\rangle = F_{\varepsilon_1}F_{\varepsilon_2}\dots F_{\varepsilon_\ell}\{s_k^0\}$$
$$n = \sum \varepsilon_j 2^{j-1}.$$

We therefore obtain a fast $2^N N$ algorithm to calculate all coefficients for "all functions in our library" for scales $-N \le j \le 0$. The procedure is analogous to subband coding.

§4. Higher dimensional libraries and bases

We define the higher dimensional wavelet basis as follows: Let

$\mathcal{W}_1 =$ span$\{W_0(x_1-k_1)W_1(x_2-k_2), W_1(x_1-k_1)W_0(x_2-k_2), W_1(x_1-k_1)W_1(x_2-k_2)\}$

$\mathcal{W}_k = \delta^k W_1$ where $\delta f(x_1 x_2) = 2f(2x_1, 2x_2)$. Clearly

$$L^2(\mathbf{R}^2) = \sum \mathcal{W}_k.$$

The basic 2-dimensional "library" consists of all functions obtained as tensor products of the one dimensional library i.e.,

$$2^{\frac{1}{2}(j_1+j_2)}W_{n_1}(2^{j_1}x_1 - k_1)W_{n_2}(2^{j_2}x_2 - k_2).$$

We will restrict our attention to the sublibraries obtained by dilating both variables by the same dilations although the case $j_2 = rj_1$ r fixed is of independent interest.

The two dimensional basis corresponding to Theorem (2.1) is given in

Theorem (4.1). *Fix ℓ, then for $(k_1, k_2) \in \mathbb{Z}^2$ $j \in \mathbb{Z}$ $2^\ell \le n < 2^{\ell+1}$,*

$$2^{j/2} W_n(2^j x_1 - k_1) 2^{\frac{j+\ell}{2}} W_0(2^{\ell+j} x_2 - k_2)$$
$$2^{\frac{(j+\ell)}{2}} W_0(2^{\ell+j} x_1 - k_1) 2^{j/2} W_n(2^j x_2 - k_2)$$
$$2^j W_{n_1}(2^j x_1 - k_1) W_{n_2}(2^j x_2 - k_2) \quad 2^\ell \le n_i < 2^{\ell+1}$$

form an orthonormal basis of L^2.

Theorem (4.2). *Let $\langle n\rangle = \langle n_1, n_2\rangle = 2^{\max(\ell(n_1)\ell(n_2))}$ where $\ell(n_1) = [\log_2 n_1]$ $n_1 \ge 0$ $n_2 \ge 0$ then for $(k_1, k_2) \in \mathbb{Z}^2$,*

$$\langle n\rangle \cdot W_{n_1}(\langle n\rangle x_1 - k_1) W_{n_2}(\langle n\rangle x_2 - k_2)$$
$$\langle 2n\rangle W_{n_1}(2\langle n\rangle x_1 - k_1) W_{n_2}(2\langle n\rangle x_2 - k_2)$$

form an orthonormal basis (wavelet packet basis of $L^2(\mathbb{R}^2)$).

Proof. Assume $\ell = \max\,(\ell(n_1)\ell(n_2)) = \ell(n_2)$, i.e. $n_1 \le n_2$. Then

$$2^\ell W_{n_1}(2^\ell x_1 - k_1) W_{n_2}(2^\ell x_2 - k_2)$$

for

$$0 \le n_1 < 2^\ell \qquad 2^\ell \le n_2 < 2^{\ell+1}$$

span (using 1.9) Proposition (1.1) $\delta^{2\ell}\Omega_{0,x_1} \otimes \delta^{2\ell}\Omega_{1,x_2}$ i.e.. the subspace spanned by $W_0(2^{2\ell}x_1 - k_1)W_1(2^{2\ell}x_2 - k_2)$. Consideration of the other 2 cases yields a total span of $\mathcal{W}_{2\ell}$, similarly we obtain $\mathcal{W}_{2\ell+1}$ and the theorem is proved.

Reference

[D] Ingrid Daubechies, *Orthonormal bases of compactly supported wavelets*, Communications on Pure and Applied Mathematics XLI (1988), 909–966.

Appendix II

Best-Adapted Wavelet-Packet Bases

Introduction. By using extra filters, it is possible to introduce fast wave packet transformations which decimate by arbitrary numbers. Such transformations generalize algorithms which decimate by 2. The method produces new libraries of orthonormal basis vectors. We introduce an algorithm for selecting a most efficient representation from the library, and prove that its complexity is $O(N)$ for a sequence of length N. We discuss some of the analytic properties and applications of such representations.

Aperiodic filters and bases in l^2. Consider first the construction of bases on l^2. Let p be a positive integer and introduce p absolutely summable sequences $f_0, \ldots, f_{p-1}$ satisfying the properties:

(1) For some $\epsilon > 0$, $\sum_m |f_i(m)||m|^\epsilon < \infty$,
(2) $\sum_m f_0(m) = 1$, for $i = 0, 1, \ldots, p-1$, and
(3) $\sum_m f_i(m) f_j(m + kp) = \delta_{i-j}\delta_k$, where δ is the Kronecker symbol.

To these sequences are associated p convolution operators $F_0, \ldots, F_{p-1}$ and their adjoints $F_0^*, \ldots, F_{p-1}^*$ defined by

$$F_i : l^2 \to l^2, \qquad F_i v(k) = \sum_m f_i(m + pk) v(m),$$
$$F_i^* : l^2 \to l^2, \qquad F_i^* v(m) = \sum_k f_i(m + pk) v(k).$$

These convolution operators will be called *filters* by analogy with quadrature mirror filters in the case $p = 2$. They have the following properties:

Lemma. *For $i, j = 0, 1, \ldots, p-1$,*

(1) $F_i F_j^* = 0, \qquad$ *if $i \neq j$,*
(2) $F_i F_i^* = I$,
(3) *$F_i^* F_i$ is an orthogonal projection of l^2, and for $i \neq j$ the ranges of $F_i^* F_i$ and $F_j^* F_j$ are orthogonal, and*
(4) $F_0^* F_0 + \cdots + F_{p-1}^* F_{p-1} = I$.

Research supported in part by ONR Grant N00014-88-K0020

Proof. Properties (1) and (2) follow by interchanging the order of integration:

$$\begin{aligned}
F_i F_j^* v(k') &= \sum_m \sum_k f_i(m+pk') f_j(m+pk) v(k) \\
&= \sum_k \left(\sum_{m'} f_i(m') f_j(m' + p[k-k']) \right) v(k) \\
&= \sum_k \delta_{i-j} \delta_{k-k'} v(k) \\
&= \begin{cases} v(k'), & \text{if } i = j, \\ 0, & \text{if } i \neq j, \end{cases}
\end{aligned}$$

by the orthogonality properties of f_i.

Property (3) follows from (1) and (2): $F_i^* F_i F_i^* F_i = F_i^* F_i$, and $F_i^* F_i F_j^* F_j = 0$. Orthogonality is easily shown by transposition.

To prove (4), let $m_j(\xi) = \sum_k f_j(k) e^{ik\xi}$ be the (bounded, Hölder continuous, periodic) function determined by the filter f_j, for $j = 0, \ldots, p-1$. Then $f_j(k) = \hat{m}_j(k)$ is a real number, and each $F_i^* F_i$ is unitarily equivalent to multiplication by $|m_i|^2$ on $L^2(-\pi, \pi)$.

Now Plancherel's theorem gives

$$\int_0^{2\pi} e^{ilp\xi} m_j(\xi) \bar{m}_{j'}(\xi)\, d\xi = \sum_k f_j(k) f_{j'}(k + lp) = \delta_{j-j'} \delta_l.$$

In particular, $|m_j|^2$ has integral 1, and the Fourier coefficient $(|m_j|^2)^\wedge(lp)$ vanishes if $l \neq 0$. This is equivalent to the average of $|m_j(\xi)|^2$ over $\{\xi, \xi + 2\pi/p, \ldots, \xi + 2\pi(p-1)/p\}$ being identically 1.

The same vanishing is true of the Fourier coefficients of the cross terms $m_j \bar{m}_{j'}$, and for those it also holds when $l = 0$. Thus, the average of $m_j(\xi) \bar{m}_{j'}(\xi)$ over $\{\xi, \xi + 2\pi/p, \ldots, \xi + 2\pi(p-1)/p\}$ vanishes identically. Hence, the conditions on the filters f_i are equivalent to the unitarity of the matrix

$$\begin{pmatrix} m_0(\xi) & m_0(\xi + \frac{2\pi}{p}) & \ldots & m_0(\xi + \frac{2\pi(p-1)}{p}) \\ \ldots & \ldots & \ldots & \ldots \\ m_{p-1}(\xi) & \ldots & \ldots & m_{p-1}(\xi + \frac{2\pi(p-1)}{p}) \end{pmatrix}$$

But then $\sum_{k=0}^{p-1} |m_k(\xi)|^2 = 1$ for all ξ. Thus $F_0^* F_0 + \cdots + F_{p-1}^* F_{p-1}$ is unitarily equivalent to multiplication by 1 in $L^2(-\pi, \pi)$, proving (4). □

With this proposition we can decompose l^2 into mutually orthogonal subspaces $W_0^1 \perp \cdots \perp W_{p-1}^1$, where $W_i^1 = F_i^* F_i(l^2)$ for $i = 0, \ldots, p-1$. The map F_i finds

the coordinates of a vector with respect to an orthonormal basis of W_i^1. Since each $F_i W_i^1 = F_i(l^2)$ is another copy of l^2, there is nothing to prevent us from reapplying the filter convolutions recursively. At the mth stage, we obtain $l^2 = W_0^m \perp \cdots \perp W_{p^m-1}^m$, where $W_n^m = F_{n_1}^* \ldots F_{n_m}^* F_{n_m} \ldots F_{n_1}(l^2)$ and $n_m \ldots n_1$ is the radix-p representation of n. The map $F_{n_m} \ldots F_{n_1}$ transforms into standard coordinates in W_n^m. For convenience, we will introduce the notations $\mathbf{F}_n^m = F_{n_m} \ldots F_{n_1}$, and $\mathbf{F}_n^{m*} = F_{n_1}^* \ldots F_{n_m}^*$.

The subspaces W_n^m form a p-ary tree. Every node W_n^m is a parent with p daughters $W_{pn}^{m+1}, \ldots, W_{pn+p-1}^{m+1}$. The root of the tree is the original space l^2, which we may label W_0^0 for consistency. Call the whole tree $\mathbf{W}$.

Now fix m and suppose w belongs to W_n^m, where $0 \leq n \leq p^m - 1$, and $\mathbf{F}_n^m w = e_k$ is the elementary sequence with 1 in the kth position and 0's elsewhere. The collection of all such w forms an orthonormal basis of l^2 with some remarkable properties. In particular, if $p = 2$ and the filters F_0 and F_1 are taken as low-pass and high-pass quadrature mirror filters, respectively, then the spaces $W_0^m, \ldots W_{2^m-1}^m$ are all the subbands at level m. These have been used for a long time in digital signal processing and compression. In an earlier paper we described experiments with an algorithm for choosing m so as to reduce the bit rate of digitized acoustic signal transmission. This produced good signal quality at rather low bit rates.

The tree contains other orthogonal bases of W_0^0. In fact, it forms a library of bases which may be adapted to classes of functions. The tree structure allows the library to be searched efficiently for the extremum of any additive functional.

To every node in $\mathbf{W}$ we associate the subtree of all its descendants. Define a *graph* to be any subset of the nodes of $\mathbf{W}$ with the property that the union of the associated subtrees is disjoint and contains a complete level $W_0^m, \ldots, W_{p^m-1}^m$ for some m. The singleton $\{W_0^0\}$ is a graph, for example, with $m = 0$. The following may be called the graph theorem.

Theorem. *Every graph corresponds to a decomposition of l^2 into a finite direct sum.*

Proof. Every graph is a finite set, of cardinality no more than p^m for the m in the definition. Fix a graph, and suppose that $W_{n_1}^{m_1}$ and $W_{n_2}^{m_2}$ are subspaces corresponding to two nodes. Without loss, suppose that $m_1 \leq m_2$. Then $W_{n_2}^{m_2}$ is contained in a subspace $W_n^{m_1}$ for some $n \neq n_1$. Since the subspaces at a given level are orthogonal, we conclude that $W_{n_2}^{m_2} \perp W_{n_1}^{m_1}$.

To show that the decomposition is complete, observe that a node contains the sum of its daughters. By induction, it contains the sum of all of the nodes in its subtree. Hence a graph contains the sum of all the subspaces at some level m. But this sum is

all of l^2. $\square$

Theorem. *Graphs are in one-to-one correspondence with finite disjoint covers of* $[0,1)$ *by p-adic intervals* $I_n^m = p^{-m}[n, n+1)$, $n = 0, 1, \ldots, p^m - 1$.

Proof. The correspondence is evidently $W_n^m \leftrightarrow I_n^m$. The subtree associated to W_n^m corresponds to all p-adic subintervals of I_n^m. The details are left to the reader. $\square$

Analytic properties of graphs: continuous wave packets. Each filter F_j (and its adjoint F_j^*) maps the class of exponentially decreasing sequences to itself. Likewise, the projections $\mathbf{F}_n^{m*}\mathbf{F}_n^m$ preserve that class. In practice, we shall consider only finite sequences in l^2. For actual computations the filters must be finitely supported as well. Convolution with such filters preserves the property of finite support. Let the support width of the filters be r, and let z_m be the maximum width of any vector of the form $F_{j_1}^* \ldots F_{j_m}^*(e_k)$. Then $z_0 = 1$ and $z_{m+1} = pz_m + r - p$. By induction, we see that $z_m = p^m + (p^m - 1)(r - p)$.

Coifman and Meyer [CM] have observed that the basis elements $\mathbf{F}_n^{m*}e_k$ are related to wave packets over $\mathbf{R}$. A slightly generalized paraphrase of their construction follows. Many of the basic facts used were proved by Daubechies in [D].

Let w be a function defined by $\hat{w}(\xi) = \prod_{j=1}^{\infty} m_0(\xi/p^j)$, where m_0 is the analytic function defined by F_0, as above. Then w has mass 1, decreases rapidly, and is Hölder continuous, as proved in [D]. If m_0 is a trigonometric polynomial of degree r, then w is supported in the interval $[-r, r]$. Arranging that w has r continuous derivatives requires m_0 with degree at most $O(r)$. See [D] for a discussion of the constant in this relation for $p = 2$. Put $w_0^0 = w$, and define the family of wave packets recursively by the formula $w_{pn+j}^{m+1}(t) = \sum_{-\infty}^{\infty} f_j(i) w_n^m(pt - i)$. This produces one function w_n^m for each pair (m, n), where $m = 0, 1, \ldots$ and $n = 0, 1, \ldots, p^m - 1$.

We can renormalize the wave packets to a fixed scale p^L. Write

$$w_{n,m,k}^L(t) = p^{(L-m)/2} w_n^m(p^{L-m}t - k).$$

Then $w_{0,0,k}^L$ is a collection of orthonormal functions of mass 1, concentrated in intervals of size $O(p^{-L})$. This makes them suitable for sampling continuous functions. Let $x(t)$ be any continuous function, and put

$$s_0^0(k) = \langle x, w_{0,0,k}^L \rangle = \int_{-\infty}^{\infty} x(t) w_0^0(p^L t - k)\, dt.$$

We may use $s_0^0(k)$ as a representative value of $x(t)$ in the interval $I_k^L = p^{-L}[k, k+1)$. The closeness of the approximation to values of x depends, of course, on the smoothness

of x. Suppose that x is Hölder continuous with exponent ϵ. Then if t_0 is any point in I_k^L, we have

$$|x(t_0) - s_0^0(k)| = |\int_{I_k^L} (x(t_0) - x(t))\, w_0^0(p^L t - k)\, dt| = O(p^{-\epsilon L}).$$

We can also take advantage of differentiability of x if we construct w_0^0 with vanishing moments. Given d vanishing moments and d derivatives of x, the approximation improves to $|x(t_0) - s_0^0(k)| = O(p^{-dL})$.

The map $x \mapsto s_0^0$ sends $L^2(\mathbf{R})$ to l^2, and pulls back the orthonormal bases of l^2 constructed in the last section. To see this, define $s_n^m(k) = \langle x, w_{n,m,k}^L \rangle$. By interchanging the order of recurrence relation and inner product, we obtain the formula $s_n^m = \mathbf{F}_n^m s_0^0$. Thus, the coordinates $s_n^m(k)$ are coefficients with respect to an orthonormal basis of W_n^m.

The resulting subspaces of $L^2(\mathbf{R})$ form a finer type of multiresolution decomposition than that of Mallat [Ma]. The coordinates $s_n^m(k)$ are rapidly computable. They contain a mixture of location and frequency information about x.

Ordering the basis elements. The parameters n, m, k, L in $w_{n,m,k}^L$ have a natural interpretation as frequency, scale, position, and resolution, respectively. However, n is not monotonic with frequency, because our construction yields wave packets in the so-called Paley (natural, or p-adic) ordering. The following results show how to permute $n \mapsto n'$ into a frequency-based ordering.

Theorem. *We can choose rapidly decreasing filters $F_0, \ldots, F_p - 1$ such that $w_{n,m,k}^L$ is concentrated near the interval I_k^{L-m}, and $\hat{w}_{n,m,k}^L$ is concentrated near the interval $I_{n'}^m$, where $n \mapsto n'$ is a permutation of the integers.*

Proof. The first part is evident. For any family of exponentially decreasing filters, w_0^0 decreases exponentially away from $[0, 1)$. $w_{0,m,k}^L$ is its dilate and translate to the interval I_k^{L-m}. Likewise, $w_{n,m,k}^L$ has the same concentration as $w_{0,m,k}^L$, since all the filters F_i are uniformly exponentially decreasing.

The second part follows from the Fourier transform of the recurrence relation:

$$\hat{w}_{pn+j}^{m+1}(\xi) = \left(p^{-1} \sum_k f_j(k) e^{-ix\xi/p}\right) \hat{w}_n^m(\xi/p) = p^{-1} m_j(\xi/p) \hat{w}_n^m(\xi/p),$$

where m_j is the multiplier defined above. Recall that $\sum_{j=0}^{p-1} |m_j(\xi)|^2 \equiv 1$ and that $m_0(0) = 1$. Thus, the periodic functions $|m_j|^2$ form a partition of unity into p functions. with 0 being in the support of m_0 alone.

Now suppose for simplicity that we have chosen filters in such a way that $|m_j(\xi)| = \sum_{k=-\infty}^{\infty} \chi_{\pm\frac{\pi}{p}[j,j+1)}(\xi - 2\pi k)$. Such m_j may be approximated in $L^2(-\pi,\pi)$ as closely as we like by multipliers arising from exponentially decreasing filters. In this simple case, it is immediate that $\hat{w}_0^0(\xi) = m_0(\xi/p)|_{(-\pi,\pi)}$ is the characteristic function of $(-\pi,\pi)$, so that $\hat{w}_{0,0,0}^L$ is the characteristic function of $(-\pi p^L, \pi p^L)$. Likewise, $\hat{w}_{j,1,0}^L$ is the characteristic function of $\pi p^{L-1}(-j-1,-j] \cup \pi p^{L-1}[j,j+1)$. From the recurrence relation, we see that $\hat{w}_{n,m,0}^L$ will be the characteristic function of the union of the intervals $\pm\pi p^{L-m}[n', n'+1)$, where $n \mapsto n'$ is a permutation. These intervals cover $p^L(-\pi,\pi)$ as $n = 0,\ldots,p^m - 1$. The permutation $n \mapsto n'$ is given by the recurrence relation

$$n' = n, \qquad \text{if } n = 0,\ldots,p-1; \qquad (np+j)' = \begin{cases} n'p + j, & \text{if } n' \text{ is even.} \\ n'p + (p-1) - j, & \text{if } n' \text{ is odd.} \end{cases}$$

Write n_j for the jth digit of n in radix p, numbering from the least significant. Set $n_m = 0$ if n has fewer than m digits. Then the recurrence relation implies that $n_j = \pi(n'_{j+1}, n'_j)$, where

$$\pi(x,y) = \begin{cases} y, & \text{if } x \text{ is even,} \\ p-1-y, & \text{if } x \text{ is odd..} \end{cases}$$

For each value of the first variable, π is a permutation of the set $\{0,\ldots,p-1\}$ in the second variable. Thus the map $n' \mapsto n$ and its inverse $n \mapsto n'$ are permutations of the integers. It is not hard to see that these are permutations of order 2 if p happens to be odd. Otherwise they have infinite order, as may be seen by considering an increasing sequence of integers n' all of which have only odd digits in radix p. □

Corollary. *With filters $F_0,\ldots,F_{p-1}$ chosen as above, we can modify the recurrence relation for $w_{n,m,k}^L$ such that $\hat{w}_{n,m,k}^L$ is concentrated near the interval I_n^m.*

Proof. Simply reorder the functions w_n^m by using the alternative recurrence relation:

$$w_{pn+j}^{m+1}(t) = \begin{cases} \sum_k f_j(k) w_n^m(pt-k), & \text{if } n \text{ is even,} \\ \sum_k f_{p-1-j}(k) w_n^m(pt-k), & \text{if } n \text{ is odd.} \end{cases}$$

Since we are enforcing $n = n'$ at each level m, we are composing with the permutation defined above. Of course, this algorithm has complexity identical to the original. □

Periodic filters and bases for $\mathbf{R}^d$. A sampled periodic function may be represented as a vector in $\mathbf{R}^d$ for some d. In this case let p be any factor of d. Introduce as filters a family of p vectors $\{\tilde{f}_i \in \mathbf{R}^d : i = 0,\ldots,p-1\}$. These are obviously summable.

Suppose in addition that they are orthogonal as periodic discrete functions, i.e., that $\sum_{m=1}^{d} \tilde{f}_i(m)\tilde{f}_j(m + kp \mod d) = \delta_{i-j}\delta_k$.

Let the associated convolution operators be $\{\tilde{F}_0, \ldots, \tilde{F}_{p-1}\}$, defined as above by

$$\tilde{F}_i : \mathbf{R}^d \to \mathbf{R}^{d/p}, \qquad \tilde{F}_i v(k) = \sum_{m=1}^{d} \tilde{f}_i(m + pk \mod d)v(m), \qquad \text{for } k = 1, 2, \ldots, d/p,$$

$$\tilde{F}_i^* : \mathbf{R}^{d/p} \to \mathbf{R}^d, \qquad \tilde{F}_i^* v(m) = \sum_{k=1}^{d/p} \tilde{f}_i(m + pk \mod d)v(k), \qquad \text{for } m = 1, 2, \ldots, d.$$

The reduction modulo d is intentionally emphasized. These operators satisfy conditions similar to those of aperiodic filters:

Proposition.

(1) $\tilde{F}_i\tilde{F}_j^* = 0$, *if* $i \neq j$,

(2) $\tilde{F}_i\tilde{F}_i^* = I_{d/p}$

(3) $\tilde{F}_i^*\tilde{F}_i$ *is a rank* d/p *orthogonal projection on* $\mathbf{R}^d$, *and for* $i \neq j$ *the ranges of* $\tilde{F}_i^*\tilde{F}_i$ *and* $\tilde{F}_j^*\tilde{F}_j$ *are orthogonal,*

(4) $\tilde{F}_0^*\tilde{F}_0 + \cdots + \tilde{F}_{p-1}^*\tilde{F}_{p-1} = I_d$

where I_d *is the identity on* $\mathbf{R}^d$.

Proof. The proof is nearly identical with the one in the aperiodic case. □

The decomposition suggested by equation (4) may be recursively applied to the p subspaces $\mathbf{R}^{d/p}$ by using additional filter families. For $d = p_1 \ldots p_L$ and $0 \leq n < d$, we have a unique representation $n = n_1 + n_2p_1 + n_3p_2p_1 + \cdots + n_Lp_{L-1} \ldots p_1$, where $0 \leq n_i < p_i$. This defines a one-to-one correspondence between $\{0, \ldots, d-1\}$ and an index set of L-tuples $I = \{(n_1, \ldots, n_L) : 0 \leq n_i < p_i\}$. We can construct a basis of $\mathbf{R}^d$ whose elements are indexed by I. For $n = (n_1, \ldots, n_L) \in I$, define $\tilde{\mathbf{F}}_n^L = \tilde{F}_{n_L}^L \ldots \tilde{F}_{n_1}^1$, where $\tilde{F}^i$ is a family of p_i periodic filters. Then $\tilde{\mathbf{F}}_n^{L*}\tilde{\mathbf{F}}_n^L$ is an orthogonal projection onto a 1-dimensional subspace of $\mathbf{R}^d$. This is shown by induction on the rank in (3). Now let W_n^L be the range of this projection. The collection $\{u_n = \tilde{\mathbf{F}}_n^{L*}1 : n \in I\}$ of standard basis vectors of W_n^L will be an orthonormal basis of $\mathbf{R}^d$, and the map $\tilde{\mathbf{F}}_n^L : \mathbf{R}^d \to \mathbf{R}$ gives the component in the u_n direction.

As before, we are not limited to the basis defined by the index set I. Products of fewer than L filters form orthogonal projections onto a tree of subspaces of $\mathbf{R}^d$. A node arising from a product of m filters will correspond to the subspace $W_n^m = \tilde{\mathbf{F}}_n^{m*}\tilde{\mathbf{F}}_n^m\mathbf{R}^d$, where $n = n_1 + \cdots + n_mp_{m-1} \cdots p_1$ indexes a composition of m filters. The tree will

be nonhomogeneous in general, although all nodes i levels from the root will have the same number p_i of daughters. Define a *nonhomogeneous graph* as a finite union of nodes whose associated subtrees form a disjoint cover of some level $m \leq L$. A graph theorem holds for this tree of subspaces as well. It and its corollary may be stated as follows:

Theorem. *Every nonhomogeneous graph corresponds to an orthogonal decomposition of* $\mathbf{R}^d$. □

Corollary. *Graphs are in one-to-one correspondence with finite disjoint covers of* $[0, 1)$ *by intervals of the form* $I_n^m = (p_1 \dots p_m)^{-1}[n, n+1)$. □

Any permutation of the prime factors of d gives a (possibly different) basis.

Smooth filters. Some filter sequences have a smoothness property:

Definition. *A summable sequence f is a smooth filter (of rank p) if there is a nonzero solution ϕ in $L^1(\mathbf{R}) \cap L^2(\mathbf{R}) \cap C^\infty(\mathbf{R})$ to the functional equation*

$$\phi(x) = p^{1/2} \sum_m f(m)\phi(px + m).$$

A filter will be said to have smoothness degree r if it satisfies this definition with C^∞ replaced by C^r. Daubechies has shown in [D] that finitely supported filters of any degree of smoothness may be constructed in the case $p = 2$. An obvious consequence is that smooth filters exist in the case $p = 2^q$. For arbitrary p, we may construct a filter family as above subject to additional constraints.

A continuous L^2 solution to the functional equation (3) always exists for a sequence f satisfying the three conditions at the top of this article. Its Fourier transform may be constructed by iteration:

$$\hat{\phi}(\xi) = \hat{\phi}(0) \prod_{i=k}^{\infty} m(\xi/p^k).$$

where $m(\xi) = p^{-1/2} \sum_k f(k)e^{-ik\xi}$ is the multiplier corresponding to the filter convolution in the integral equation. If ϕ is nonzero, then $\phi(0) \neq 0$. so it may be assumed that $\phi(0) = 1$. Now the sequence $\{f(k)|k|^\epsilon\}$ converges absolutely, so $m(\xi)$ is Hölder continuous of degree ϵ. But also, $m(0) = p^{1/2} \sum_k f(k) = 1$. so that for ξ near 0 the estimate $|m(\xi) - 1| < C|\xi|^\epsilon$ holds. This implies that the infinite product converges.